

ESTADÍSTICA DESCRIPTIVA E INDICADORES ECONÓMICOS

Por

Myladis Cogollo Flórez

Magister en Ciencia Estadística

Universidad EAFIT

Facultad de Ciencias y Humanidades

Departamento de Ciencias Básicas

Noviembre de 2008

Contenido

1. Introducción	5
1.1. Algunos conceptos básicos	5
2. Análisis de la información de variables categóricas	8
2.1. Distribución de frecuencias	8
2.2. Representación gráfica	9
2.2.1. Diagrama de barras	11
2.2.2. Diagrama circular	14
3. Análisis de la información de variables cuantitativas	14
3.1. Distribución de frecuencias	14
3.1.1. Tabla de frecuencias para datos agrupados	17
3.2. Representación gráfica	20
3.2.1. Histograma de frecuencias	20
3.2.2. Polígono de frecuencias	22
3.2.3. Ojiva	23
3.2.4. Diagrama de cajas o boxplot	24
4. Indicadores	27
4.1. Razón y/o índice, Tasa y Proporción	27
4.2. Definición e importancia de un indicador	29
4.2.1. Tipos de indicadores	29
4.3. Pirámides e Índices demográficos	30
4.3.1. Pirámide poblacional	30
4.3.2. Indicadores demográficos	31
4.4. Indicadores de calidad de vida	33
4.4.1. Indicadores de Impacto	33
4.4.2. Indicadores sectoriales	34

Índice de Figuras

1.	Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al gran total)	12
2.	Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al gran total)	12
3.	Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al total de columna)	13
4.	Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al total de fila)	13
5.	Diagrama de barras para el número de cilindros	16
6.	Diagrama de barras para el número de cilindros según el origen del automóvil . .	17
7.	Histograma de los diámetros de los tubos.	21
8.	Tipología de las distribuciones de frecuencias agrupadas.	22
9.	Polígono de frecuencias de los diámetros de los tubos.	23
10.	Ojiva para los diámetros de los tubos.	24
11.	Diagrama de cajas de los diámetros de los tubos.	26
12.	Casos mensuales nuevos de malaria en Antioquía (1998-2002).	26
13.	Ilustración de la limitación de los diagramas de cajas.	27

Índice de Tablas

1.	Tabla de frecuencias para las variables categóricas	8
2.	Frecuencias. Encuesta sobre el hábito de fumar	9
3.	Clasificación de los empleados de una empresa, según estado civil y género	9
4.	Clasificación de los empleados de una empresa, según estado civil y género (Porcentaje con respecto al gran total)	10
5.	Clasificación de los empleados de una empresa, según estado civil y género (Porcentaje con respecto al total de columna)	10
6.	Clasificación de los empleados de una empresa, según estado civil y género (Porcentaje con respecto al total de fila)	10
7.	Tabla de frecuencias del número de cilindros	15

8.	Estadísticos básicos para la variable número de cilindros	16
9.	Tabla de frecuencias para datos agrupados	19
10.	Tabla de frecuencias para los diámetros de 30 tubos	20

1. Introducción

La Estadística es la Ciencia de la sistematización, toma, ordenación y presentación de los datos referentes a un fenómeno que presenta variabilidad o incertidumbre para su estudio metódico, con objeto de deducir las leyes que rigen esos fenómenos, y poder de esa forma hacer previsiones sobre los mismos, tomar decisiones u obtener conclusiones. El campo de aplicación de la estadística es muy amplio, se aplica en la industria, en la medicina, en las finanzas, en la investigación de mercados, entre otros.

1.1. Algunos conceptos básicos

Población: es el conjunto de individuos o elementos, sobre el que estamos interesados en obtener conclusiones (hacer inferencia). Según el tamaño, la población puede considerarse finita ó infinita. Es de resaltar el hecho de que a pesar de encontrarse una población constituida por un grupo d elementos, a la estadística no le interesa el elemento en sí, sino sus características.

Experimento: Uno de los objetivos de la Estadística es el poder obtener conclusiones a partir de la información recolectada de una población de interés. Esta información es usualmente obtenida a partir de un experimento.

Un experimento es cualquier procedimiento que genere información. Se pueden identificar dos tipos de experimentos:

- **Experimento determinístico:** Si las condiciones bajo las cuales se realiza el experimento determinan completamente los resultados del mismo. Por ejemplo, al lanzar una moneda cargada, la forma de maracr el número 2619500, el ancho de un cuarto.
- **Experimento aleatorio:** Cuando sus resultado no puede ser detrmnado de antemano, aunque puede conocerse el conjunto de todos los posibles resultados. Por ejemplo, la presión arterial de una persona puede variar en el transcurso del día o diariamente, el tiempo entre llamadas de un celular cambia de llamada en llamada, diámetro de las tuercas.

Muestra: es un subconjunto de la población, al que tenemos acceso (a través de un experimento) y sobre el que realmente hacemos las observaciones (mediciones). La muetsra será *representativa* si las probabilidades de selección de los elementos son todas iguales, durante todo

el proceso, y se denominará *Muestra aleatoria*. Los conceptos básicos para delimitar una muestra aleatoria correctamente son:

- **Elemento/unidad muestral:** cada uno de los orígenes de la información/componentes de la muestra.
- **Marco Muestral:** También conocido como marco de muestreo, es un dispositivo que permite identificar y ubicar a cada uno de los elementos de la población. Los tres errores más graves de un marco, los cuales deben ser revisados y corregidos a fin de evitar sesgos en las estimaciones, son:
 1. *Subcobertura:* Todos los elementos de la población deben estar en el marco. Esto implica que en el mantenimiento del marco se disponga de un mecanismo para registrar todo nuevo nacimiento o aparición de nuevos elementos.
 2. *Sobrecobertura:* Todos los elementos del marco deben ser elementos de interés en la población. Esto quiere decir que en el mantenimiento del marco se dispone de un mecanismo de registro de bajas.
 3. *Repetición:* Los elementos de la población que se encuentren registrados deben estar una y sólo una vez en el marco.
- **Alcance de la muestra:** limitaciones establecidas por el investigador en la descripción del marco muestral.
- **Parámetro:** Valor poblacional desconocido.
- **Estimador o Estadístico:** Cualquier función de la muestra totalmente especificada, es decir, que no dependa de parámetros desconocidos. El valor obtenido al evaluar la función en los valores de la muestra se denomina estimación del parámetro.
- **Error aleatorio de muestreo, error de muestreo o error aleatorio:** Diferencia entre los resultados obtenidos de una muestra y los obtenidos sobre la población, habiendo utilizado procedimientos idénticos para obtener la información.

En realidad, la diferencia no es medible (¿por qué?), por ello el investigador someterá esa diferencia a que no sobrepase un determinado límite. *A menos error muestral, el resultado es más preciso.*
- **Error sistemático, ajeno al muestreo o no muestral:** No está relacionado con la selección de la muestra. Puede ser consecuencia de la naturaleza del diseño del estudio

y de las imperfecciones en su ejecución.

Este tipo de error NO PUEDE SER MEDIDO, por lo cual debe evitarse al máximo.

Variable: es una característica observable que varía entre los diferentes individuos de una población. La información que disponemos de cada individuo es resumida en variables. Las variables pueden ser de dos tipos:

- **Variables cualitativas o categóricas:** Las variables cualitativas son aquellas que no aparecen en forma numérica, sino como categorías o atributos como, por ejemplo, el género o la profesión de una persona. En dichas categorías puede haber un orden subyacente (*variable ordinal*) o no (*variable nominal*). Los datos correspondientes a variables cualitativas se agrupan de manera natural en diferentes categorías o clases y se cuenta el número de datos que aparecen en cada una de ellas (no se pueden hacer operaciones algebraicas con ellos).
- **Variables cuantitativas:** Las variables cuantitativas son las que pueden expresarse numéricamente. Una primera clasificación, basada en el tipo de valores que puede tomar, permite distinguir entre *variables cuantitativas discretas*, que son, frecuentemente el resultado de contar y, por tanto, toman sólo valores enteros y *continuas*, que resultan de medir y pueden contener cifras decimales. Variables discretas son el número de lavadoras producidas por una empresa en un año, número de hijos, número de cigarrillos que se fuma diariamente, etc. Variables continuas son aquellas cuyos valores pueden ser cualquier cantidad en un intervalo, como la temperatura, el peso o la altura de una persona, la superficie de las viviendas, el tiempo. Además, las expresadas en medidas de relación, tales como porcentajes, tasas, puntuaciones.

■ **Ejemplo 1.** *Una máquina produce piezas que se pueden clasificar en defectuosas y válidas; se quiere saber cuál es la proporción de piezas defectuosas que produce la máquina.*

Población: *Infinitas piezas que produce la máquina.*

Variable de interés: *Número de piezas defectuosas.*

Muestra: *Las piezas seleccionadas para comprobar si son defectuosas.*

Unidad muestral: *Piezas producidas.*

Marco muestral: *Registro de las piezas producidas por la máquina.*

Problema: *No conocemos el porcentaje de piezas defectuosas (p). Nótese que en este caso, p es el parámetro. Queremos hacer inferencias acerca del verdadero valor de p .*

NOTA: *Nunca se sabe el valor real de p . Sólo es posible estimarlo.*

2. Análisis de la información de variables categóricas

2.1. Distribución de frecuencias

Las tablas de frecuencias corresponden a una agrupación o clasificación de los datos, con el fin de determinar la frecuencia con que se repite el atributo o el valor que toma la variable.

Con respecto a las variables cualitativas se ha dicho que no aparecen de manera numérica sino como atributos. Generalmente, los atributos son dicotómicos, de modo que cada miembro de la población se incluye en uno de los dos grupos conforme posea una característica específica o no la posea (Por ejemplo: Fuma o no fuma). Aunque los atributos también permiten clasificaciones múltiples, ya que una persona puede tener preferencia por más una opción, por ejemplo, sobre una marca de un producto determinado. La tabulación en atributos es muy sencilla, pues su cuantificación se hace mediante el conteo del número de veces que se presenta o se repite la característica investigada. La estructura general de la tabla de frecuencias, para una variable con k modalidades o categorías, es la siguiente: donde,

Tabla 1: Tabla de frecuencias para las variables categóricas

Nombre de la variable	Frecuencia absoluta	Frecuencia relativa	Porcentaje (%)
Modalidad 1	n_1	n_1/n	$(n_1/n) \times 100$
Modalidad 2	n_2	n_2/n	$(n_2/n) \times 100$
\vdots	\vdots	\vdots	\vdots
Modalidad k	n_k	n_k/n	$(n_k/n) \times 100$
Total	n	1	100

- Frecuencias absolutas: Contabilizan el número de individuos de cada modalidad.
- Frecuencias relativas (porcentajes): Idem, pero dividido por el total

- Frecuencias acumuladas: Sólo tienen sentido para variables ordinales y numéricas. Es opcional.

■ **Ejemplo 2.** *Suponga una encuesta realizada a 20 trabajadores sobre el hábito de fumar, con los siguientes resultados:*

Tabla 2: Frecuencias. Encuesta sobre el hábito de fumar

Hábito del cigarrillo	Número de fumadores	Porcentaje %
Sí	12	60
No	8	40
Total	20	100

■ **Ejemplo 3.** *Pueden realizarse tablas de doble entrada, las cuales se denominan tablas de contingencia, por ejemplo: En algunos casos los valores que contiene la tabla de contingencia son*

Tabla 3: Clasificación de los empleados de una empresa, según estado civil y género

ESTADO CIVIL	GÉNERO		TOTAL
	Masculino	Femenino	
Soltero	28	22	50
Casado	52	10	62
Separado	17	5	22
Viudo	3	1	4
Unión libre	10	2	12
Total	110	40	150

porcentajes; y aquí surgen tres tipos de tablas (se escoge una de acuerdo a los requerimientos del estudio): Con respecto al total de datos, con respecto al total de fila o con respecto al total de columna. Para el ejemplo tenemos:

2.2. Representación gráfica

Se utiliza para visualizar mejor la información y como complemento de la tabla de frecuencias. Las variables categóricas se suelen representar mediante diagrama de barras, sectores o pastel.

Tabla 4: Clasificación de los empleados de una empresa, según estado civil y género (Porcentaje con respecto al gran total)

ESTADO CIVIL	GÉNERO		TOTAL
	Masculino	Femenino	
Soltero	18.7	14.7	33.3
Casado	34.7	6.7	41.3
Separado	11.3	3.3	14.7
Viudo	2	0.7	2.7
Unión libre	6.7	1.3	8.0
Total	73.3	26.7	100

Tabla 5: Clasificación de los empleados de una empresa, según estado civil y género (Porcentaje con respecto al total de columna)

ESTADO CIVIL	GÉNERO	
	Masculino	Femenino
Soltero	25.5	55.0
Casado	27.3	25.0
Separado	15.5	12.5
Viudo	2.7	2.5
Unión libre	9.1	5.0
Total	100	100

Tabla 6: Clasificación de los empleados de una empresa, según estado civil y género (Porcentaje con respecto al total de fila)

ESTADO CIVIL	GÉNERO		TOTAL
	Masculino	Femenino	
Soltero	56.0	44.0	100
Casado	83.9	16.1	100
Separado	77.3	22.7	100
Viudo	75.0	25.0	100
Unión libre	83.3	16.7	100

2.2.1. Diagrama de barras

Cada modalidad de la variable se representa por un rectángulo con altura proporcional a la frecuencia absoluta o al porcentaje correspondiente. Todos los rectángulos deben ser del mismo ancho y se debe procurar que el espacio entre las barras sea la mitad del ancho de la barra. Dos o tres grupos de información pueden ser comparados por grupos de barras. Las barras pueden trazarse en forma horizontal o vertical. Además, las partes componentes de un total pueden representarse subdividiendo las alturas de cada rectángulo. Este tipo de gráfico es especial para representar distribuciones combinadas, donde una variable es cuantitativa y la otra cualitativa.

■ **Ejemplo 4.** A continuación se muestra la distribución de frecuencias del Nivel de Estudios de cierta población. Así mismo, se muestra el diagrama de barras correspondiente a dicha variable. El nivel de estudios más frecuente es la categoría Sin Estudios, mientras que la categoría mediana

Tabla 3. Tabla de frecuencias del Nivel de Estudios

		Nivel de estudios			
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Sin estudios	217199	36.1	36.3	36.3
	Bachiller elemental	199625	33.2	33.4	69.7
	Bachiller superior	104726	17.4	17.5	87.2
	Diplomado	36573	6.1	6.1	93.3
	Licenciado	40261	6.7	6.7	100.0
	Total	598384	99.5	100.0	
Perdidos	Sistema	3288	.5		
Total		601672	100.0		

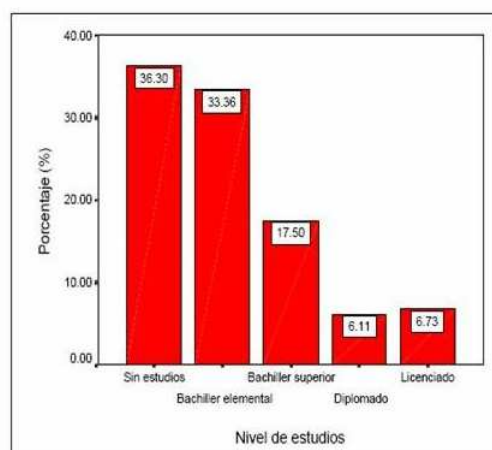


Figura 1: Diagrama de Barras del Nivel de Estudios

es Bachiller Elemental. Se observa así mismo que, tan sólo un 12.84 % tienen estudios superiores poniéndose de manifiesto, además, que existen más licenciados que diplomados en la ciudad debido, muy probablemente, al escaso número de Escuelas Universitarias existentes en la misma.

■ **Ejemplo 5.** Para las tablas de contingencia mostradas en el ejemplo 3, se tienen los siguientes gráficos:

Inicialmente, nótese que las figuras 1 y 2 son equivalentes; a partir de cualquiera de ellas, se deduce una cierta tendencia en el comportamiento e los géneros dentro de cada estado civil, siempre hay más hombres que mujeres. Esto se posiblemente ocurre porque en general en la en-

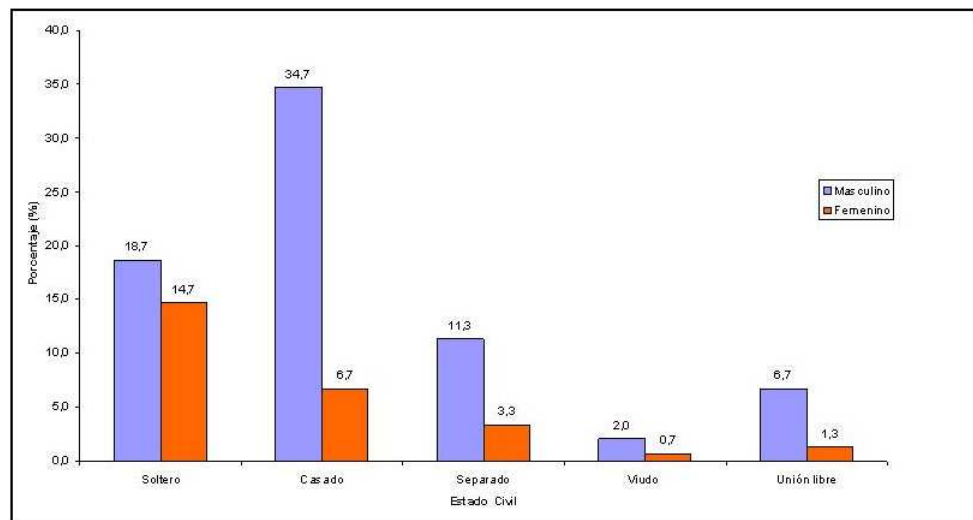


Figura 1: Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al gran total)

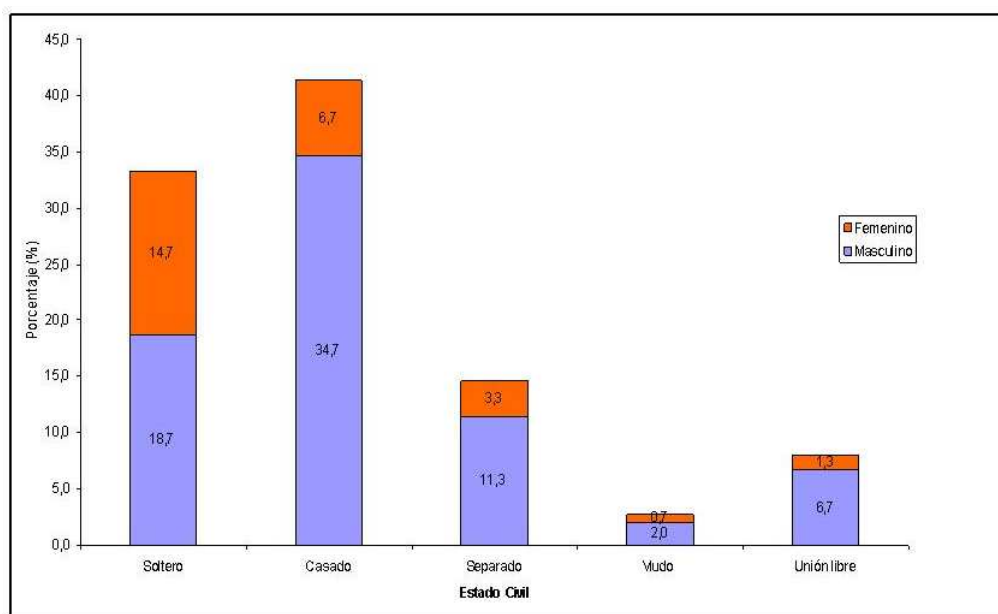


Figura 2: Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al gran total)

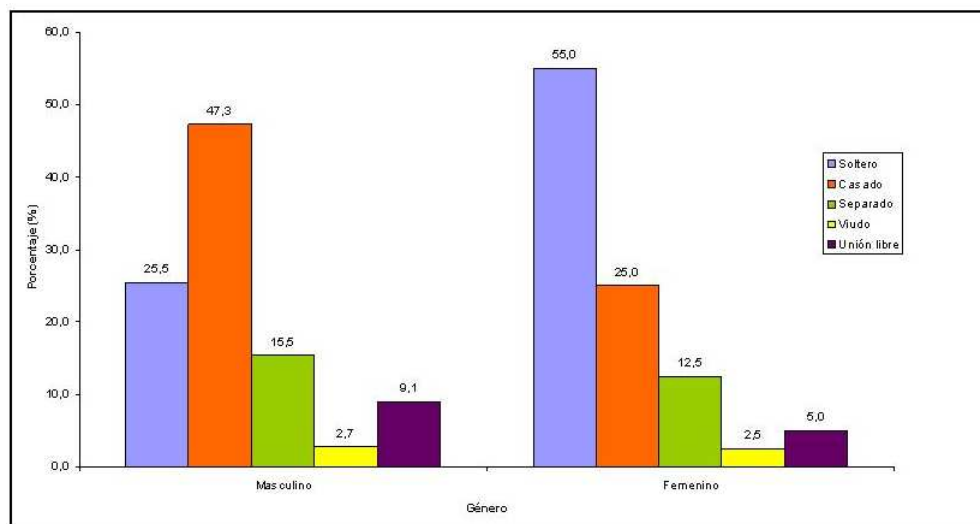


Figura 3: Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al total de columna)

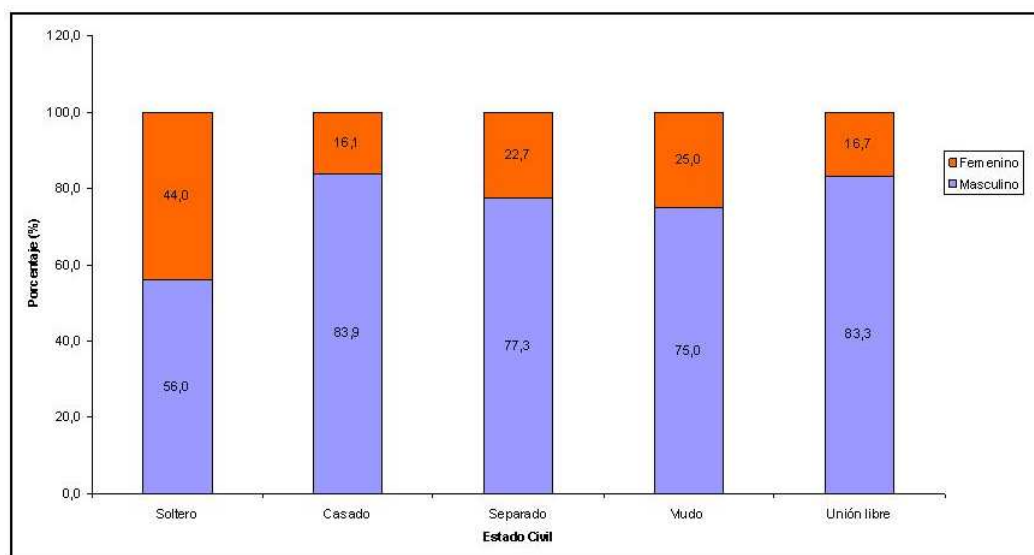


Figura 4: Diagrama de barras para la clasificación de los empleados según estado civil y género (Porcentaje con respecto al total de fila)

cuesta participaron más hombres (110) que mujeres(40). masculino. Además, se observa que en la empresa el 34.7 % de los empleados son hombres casados.

A partir de la figura 3 se tiene que en la población de empleados hombres de la empresa, predominan los de estado civil casado (47.3 %), seguidos de los solteros (25.5 %). Sin embargo, en la población de mujeres, predominan las solteras (55 %), seguidas de las casadas (25 %). Para el resto de estados civiles el comportamiento es muy similar en las dos poblaciones.

En la figura 4, se concluye nuevamente que dentro de cada estado civil siempre hay una mayor frecuencia de hombres que de mujeres.

2.2.2. Diagrama circular

Denominado por su forma pastel o torta, es otro tipo de gráfico que permite observar los componentes de un total, como sectores de un círculo. Es utilizado en las variables cualitativas para hacer notar las diferencias en las proporciones o porcentajes en que se presenta un hecho de acuerdo con sus características; consiste en subdividir los 360° de un círculo, proporcionalmente al número o al porcentaje de cada una de las modalidades de la variable.

■ **Ejemplo 6.** *La tabla de frecuencias corresponde al Estado Civil de una muestra extraída de los clientes de un Supermercado. Así mismo, en la Figura se muestra el diagrama de sectores correspondiente a esta variable. Observe que la mayor parte de los clientes (75.9 %) son casados que constituye el valor modal de la distribución de frecuencias, y que apenas acuden personas separadas (4 %).*

3. Análisis de la información de variables cuantitativas

3.1. Distribución de frecuencias

Las variables cuantitativas discretas con un número pequeño de valores se tratarían de manera similar a las variables cualitativas antes descritas.

■ **Ejemplo 7.** *En la Tabla 7 se muestra la distribución de frecuencias del Número de cilindros que tienen los 90 automóviles, de tres orígenes (Americanos, Europeos y Japoneses), que fueron puestos a prueba para calificar su rendimiento de acuerdo a las millas por galón, al recorrido total, a la potencia, al precio, entre otras variables. Así mismo, la Tabla 8 presenta algunas*

Tabla 4. Tabla de frecuencias del Estado Civil

Estado Civil		
	Frecuencia	Porcentaje
Soltero	77	19.2
Casado	305	75.9
Viudo	16	4.0
Separado	4	1.0
Total	402	100.0

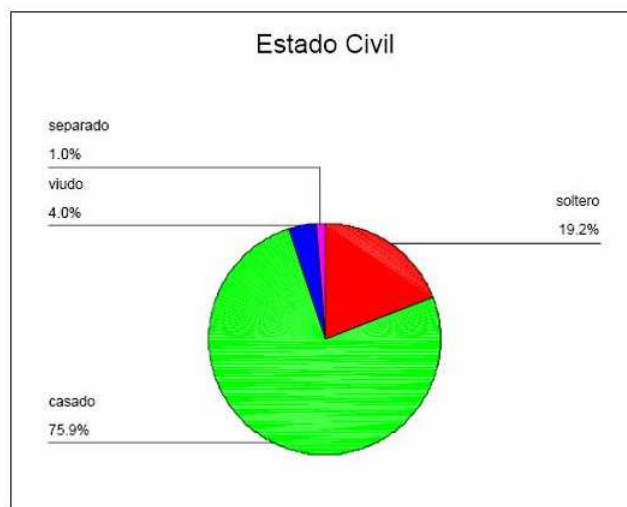


Figura 2 : Diagrama de Sectores del Estado Civil

medidas descriptivas numéricas de dicha distribución y las Figuras 5 y 6 los diagramas de barras para el número de cilindros y el número de cilindros por origen, respectivamente.

Tabla 7: Tabla de frecuencias del número de cilindros

Número de cilindros	Frecuencia	Porcentaje %
3	2	2.2
4	50	55.6
5	1	1.1
6	27	30.0
8	10	11.1
Total	88	100

El número medio de cilindros que tienen los automóviles es 4, siendo éste el valor más frecuente y el valor mediano. En el 50 % de los casos el número de cilindros oscila entre 4 y 6 (ver Tabla 7).

Note que la mayoría de los automóviles del experimento tenían 4 cilindros, independientemente del origen (ver Figuras 5 y 6).

Si la variable analizada es continua o discreta con un elevado número de valores distintos se tabula como una distribución de frecuencias agrupadas y se representa gráficamente mediante

Tabla 8: Estadísticos básicos para la variable número de cilindros

Estadísticos	Estimación	
N	Válidos	90
Media		5.03
Mediana		4.0
Moda		4.0
Des. Típ.		1.4018
Coef. Varia.		27.8512 %
Asimetría		0.8906
Curtosis		-0.3358
Mínimo		3
Máximo		8
Percentiles	25	4
	50	4.0
	75	6

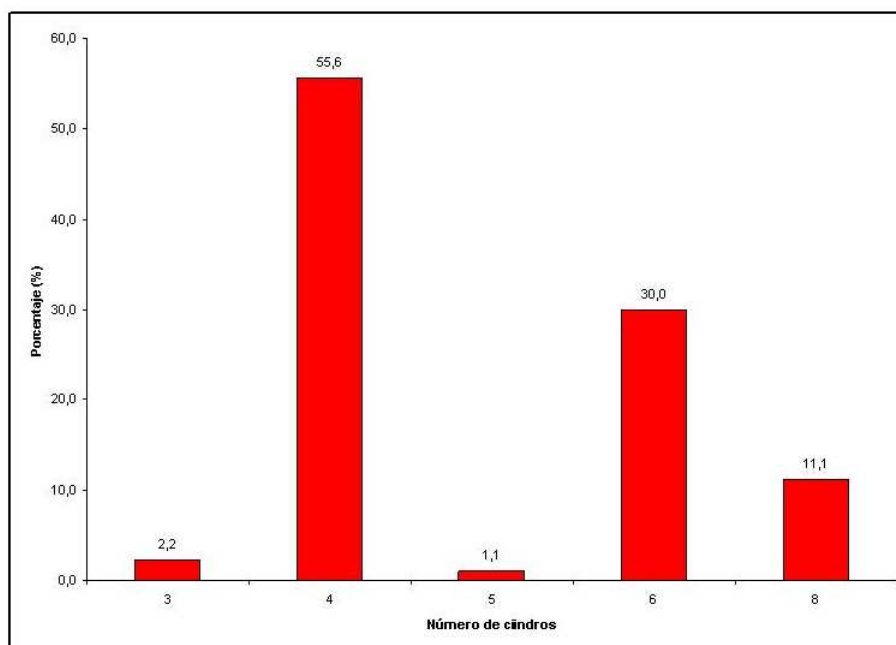


Figura 5: Diagrama de barras para el número de cilindros

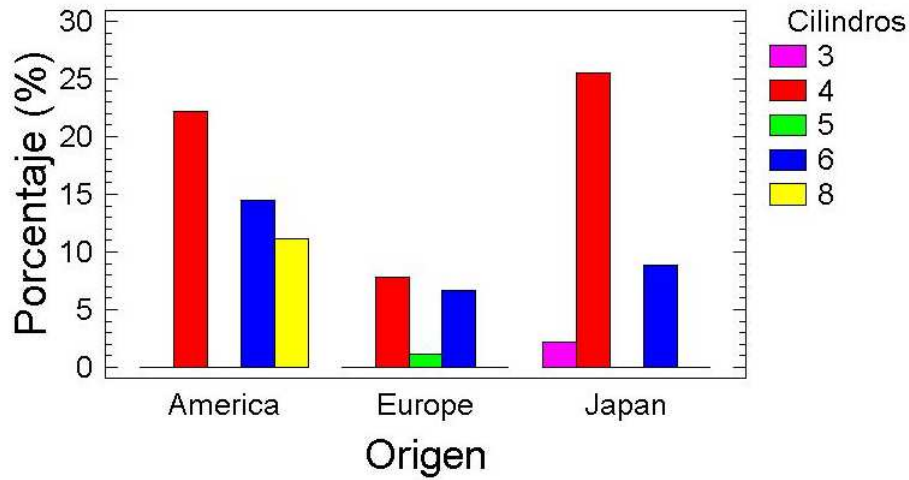


Figura 6: Diagrama de barras para el número de cilindros según el origen del automóvil

histogramas, polígonos de frecuencias, ojivas y boxplots con el fin de estudiar la forma de la distribución y analizar, en particular, la posible existencia de varias modas en la misma que pongan de manifiesto la presencia de diversos grupos homogéneos en la muestra. En la próxima se definen las tablas de frecuencias para datos agrupados y las representaciones gráficas mencionadas anteriormente.

3.1.1. Tabla de frecuencias para datos agrupados

Los pasos a seguir para construir una tabla de frecuencias agrupando las observaciones de una variable X , son:

1. *Determinar el valor máximo (X_{max}) y mínimo (X_{min}) que toma X .*
2. *Calcular el rango de X : $R = X_{max} - X_{min}$.*
3. *Determinar el número de clases (k):*

Se recomienda usar entre 5 y 16 clases. Una sugerencia útil para hallar k es considerar la regla de Sturges:

$$k = 1 + 3.322 \log(n)$$

Nota: Usualmente k no es un número entero y hay que aproximar matemáticamente.

4. *Calcular la amplitud de clase (h):*

Como k clases deben cubrir el rango R , entonces éste se divide entre el número de clases para encontrar la amplitud de clase:

$$h = \frac{R}{k}$$

Nota: Generalmente h no es de la misma exactitud de los datos. en estos casos se debe aproximar h al número más próximo (valga la redundancia) por encima que tenga la misma exactitud de los datos.

5. *Determinar el nuevo rango (R^*):*

Para determinar en qué valor empieza la primera clase se tiene en cuenta si al escoger la amplitud h , el rango inicial se amplía o no, para ello se calcula el nuevo rango $R^* = h \times k$ y se consideran los siguientes casos:

Caso I: Si el rango no se amplía, esto es, $R^* - R = 0$, entonces la primera clase empieza en el valor mínimo de la muestra. En este caso, Nuevo mínimo: $X_{min}^* = X_{min}$.

caso II: Si el rango se amplía, es decir $R^* - R = a > 0$ se distribuye convenientemente las unidades en las que se amplía: Nuevo mínimo: $X_{min}^* = X_{min} - (a/2)$ y Nuevo máximo: $X_{max}^* = X_{max} + (a/2)$.

6. *Determinar los límites reales de clases (LRC):*

Con los valores X_{min}^* y X_{max}^* se construye cada intervalo de clase:

$$X_1 - X_2, X_2 - X_3, \dots, X_{k-1} - X_k$$

En este caso, $X_1 = X_{min}^*$ y $X_k = X_{max}^*$. Además, para evitar que los intervalos se traslapen, se hallan los límites reales de clases de la siguiente manera:

a) Si los datos son número enteros:

$$\text{Límite inferior real: } L_i = X_{i-1} + 0.1$$

$$\text{Límite superior real: } L_s = X_i + 0.1$$

b) Si los datos son número con un decimal:

$$\text{Límite inferior real: } L_i = X_{i-1} + 0.01$$

$$\text{Límite superior real: } L_s = X_i + 0.01$$

Note que X_{min}^* y X_{max}^* no se modifican.

7. Calcular las marcas de clases (m_i):

La marca de clase es el punto medio del intervalo de clase:

$$m_i = \frac{L_i + L_s}{2}$$

8. Determinar las frecuencias absolutas, relativas y las acumuladas de manera análoga a los datos sin agrupar. Sin embargo, note que en este caso n_i es el número de datos que hay en cada clase.

La tabla de frecuencias se expresa de la siguiente manera:

Tabla 9: Tabla de frecuencias para datos agrupados

Clases	LRC	m_i	frecuencia absoluta(n_i)	frecuencia relativa(f_i)	frecuencia acumulada(N_i)	porcentaje (%)
$X_1 - X_2$	$[L_{i1}, L_{s1})$	m_1	n_1	n_1/n	n_1	$(n_1/n) \times 100$
$X_2 - X_3$	$[L_{i2}, L_{s2})$	m_2	n_2	n_2/n	$n_1 + n_2$	$(n_1/n) \times 100$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$X_{k-1} - X_k$	$[L_{ik}, L_{sk})$	m_k	n_k	n_k/n	$n_1 + n_2 + \cdots + n_k$	$(n_k/n) \times 100$
Total			n	1		100

■ **Ejemplo 8.** Los siguientes datos representan el diámetro interno en cm de 30 tubos para acueducto tomados como muestra dentro de un programa de calidad estatal.

14.1	14.2	13.9	14.7	12.4	14.2	14.2	14.7	13.9	13
15	14.1	14.5	14.9	13.6	14.6	14.1	14	14.8	14.7
14.5	14	15.1	14.7	13.8	15.2	13.5	14.2	14.8	14.5

Note que la variable de interés en este problema es el diámetro interno de los tubos, la cual es una variable cuantitativa continua. Debido a que esta variable presenta un elevado número de valores distintos, entonces para construir una tabla de frecuencias es necesario agrupar los datos en intervalos. Usando Excel se encuentran los siguientes valores y con base en ellos se construye la tabla de frecuencias de la Tabla 10.

$n = 30$	$h = 0.4667 \approx 0.5$
$k = 5.9070 \approx 6$	$R^* = 6 \times 0.5 = 3$
$X_{min} = 12.4$	$a = R^* - R = 0.2$
$X_{max} = 15.2$	$X_{min}^* = 12.4 - 0.1 = 12.3$
$R = 2.8$	$X_{max}^* = 15.2 + 0.1 = 15.3$

Tabla 10: Tabla de frecuencias para los diámetros de 30 tubos

Clases	LRC	m_i	n_i	f_i	N_i	%
12.3-12.8	[12.3, 12.81)	12.56	1	0.0333	1	3.3333
12.8- 13.3	[12.81, 13.31)	13.06	1	0.0333	2	3.3333
13.3-13.8	[13.31, 13.81)	13.56	3	0.1000	5	10.0000
13.8-14.3	[13.81, 14.31)	14.06	11	0.3667	16	36.6667
14.3-14.8	[14.31, 14.81)	14.56	10	0.3333	26	33.3333
14.8-15.3	[14.81, 15.3)	15.06	4	0.1333	30	13.3333
Total			30	1		100,0000

3.2. Representación gráfica

Los siguientes gráficos se construyen a partir de la tabla de frecuencias para datos agrupados, en el plano cartesiano.

3.2.1. Histograma de frecuencias

Los pasos a seguir para construir el histograma son:

1. Ubique en el eje horizontal los límites reales de clases.
2. Ubique en el eje vertical las frecuencias relativas.
3. De cada clase levante un rectángulo de altura proporcional a la frecuencia relativa correspondiente y de base igual a la amplitud de clase.

■ **Ejemplo 9.** *A partir de la Tabla 10 se obtiene el histograma para los diámetros de los 30 tubos.*

En la Figura 7, se observa que el intervalo modal es (13.8, 14.3), además se nota que la mayoría de los tubos inspeccionados tienen diámetros superiores a 13.8. Hay muy pocos tubos

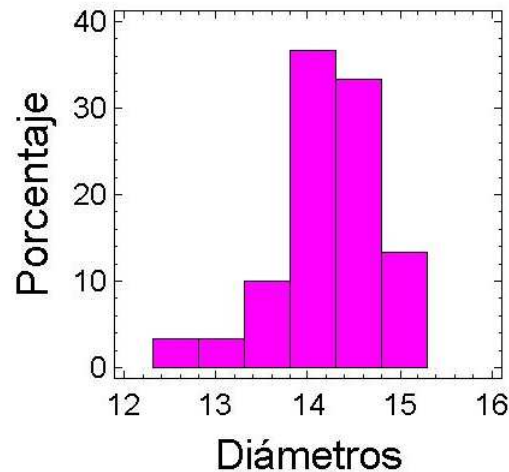


Figura 7: Histograma de los diámetros de los tubos.

con diámetros pequeños.

El hecho de que la gran mayoría de los diámetros de los tubos sean grandes, hace que en la gráfica los valores se acumulen hacia la derecha y se presente una cola (representando los pocos valores pequeños) hacia la izquierda; Esta forma del gráfico permite concluir que el diámetro de los tubos tiene una distribución unimodal asimétrica a la izquierda (o negativa).

La conclusión que se extrajo de la Figura 7, es una de las utilidades que tiene el histograma en el análisis de la forma de la distribución de una variable. La Figura 8 presenta de forma estilizada, algunos de los patrones que más frecuentemente presentan los histogramas. Las distribuciones (a) y (b) son simétricas alrededor de un valor central. El caso (a) presenta un único máximo, se dice que es una distribución unimodal, que necesariamente ha de coincidir con el centro de simetría y, en este caso, las medidas de tendencia central son una síntesis adecuada de la información contenida en la variable. La distribución (b) tiene dos máximos o modas, uno a cada lado del centro de simetría. Este patrón aparece cuando los datos responden a una mezcla de dos grupos heterogéneos y, siempre que sea posible, conviene estudiar ambos grupos por separado. Las formas que aparecen en (c) y (d) se denominan asimétrica a la derecha (o positiva) y a la izquierda (o negativa), respectivamente, e indican la presencia de un número significativo de valores muy altos (c) y bajos (d) susceptibles de distorsionar los resultados de análisis estadísticos posteriores. En estos dos últimos casos, los coeficientes de asimetría son significativamente distintos de cero, positivo en el caso (c) y negativo en el caso (d). Conviene hacer notar finalmente, que aunque

una distribución sea unimodal, no está libre de la presencia de valores anormalmente altos y bajos en ambas colas de la distribución que puedan distorsionar los resultados de un análisis estadístico. Para detectar este hecho se utiliza el coeficiente de curtosis, de forma que si la distribución es leptocúrtica (curtosis muy elevada), indica que sus colas son *muy pesadas* y, por lo tanto, se corre el riesgo antes nombrado.

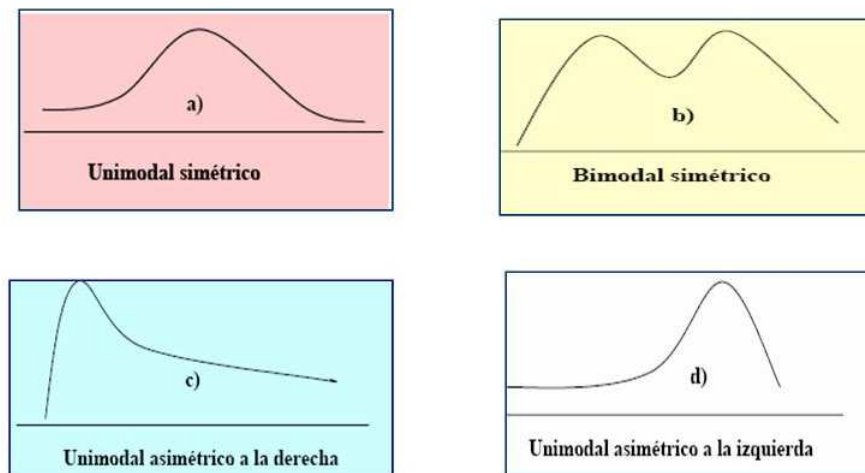


Figura 8: Tipología de las distribuciones de frecuencias agrupadas.

3.2.2. Polígono de frecuencias

El polígono de frecuencias puede realizarse de dos formas:

1. *Sobre el histograma de frecuencias*: Se unen los centros de las bases superiores de los rectángulos.
2. *A parte*:
 - Ubique en el eje horizontal las marcas de clases.
 - Ubique en el eje vertical las frecuencias relativas.
 - De cada marca de clase levante un punto de altura proporcional a la frecuencia relativa correspondiente.
 - Una los puntos con una línea continua, cerrándose en la marca de clase anterior al primer intervalo y posterior al último.

■ **Ejemplo 10.** A partir de la Tabla 10 ó del histograma de la figura 7, se obtiene el polígono de frecuencias

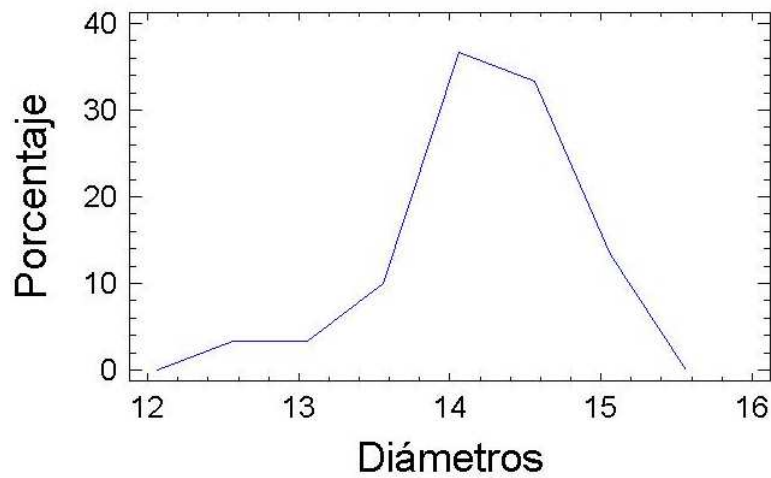


Figura 9: Polígono de frecuencias de los diámetros de los tubos.

La interpretación es similar a la de la Figura 7.

3.2.3. Ojiva

Se emplea un polígono de frecuencias con una característica particular: muestra las frecuencias absolutas (o relativas) acumuladas.

- Ubique en el eje horizontal los límites reales de clases.
- Ubique en el eje vertical las frecuencias absolutas (o relativas) acumuladas.
- De cada marca de clase levante un punto de altura proporcional a la frecuencia relativa correspondiente.
- Trace un punto sobre cada límite superior de altura proporcional a su frecuencia acumulada.

■ **Ejemplo 11.** Para los diámetros de los tubos tenemos:

Como esta gráfica presenta las frecuencias acumuladas es de gran importancia en la estadística. Por ejemplo, ¿qué porcentaje de tubos tiene un diámetro superior a 14 cm? A partir de la ojiva, se concluye que la proporción de tubos equivale aproximadamente a un 70 %.

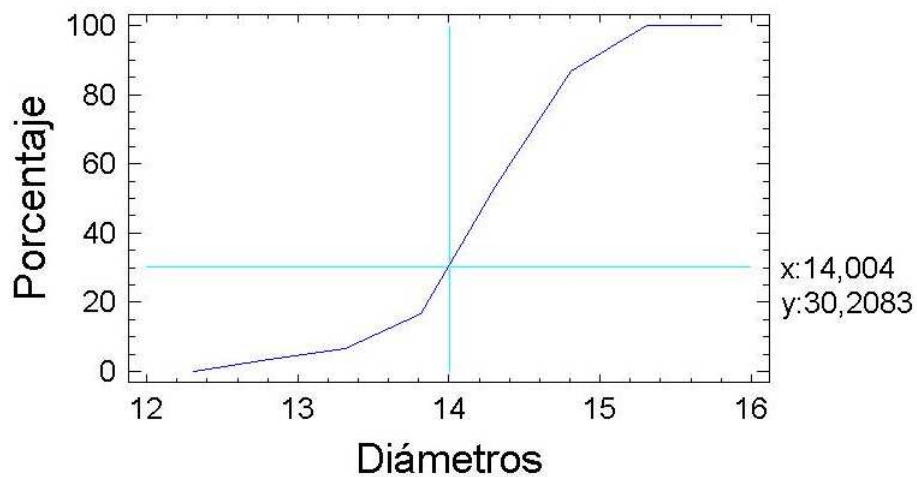


Figura 10: Ojiva para los diámetros de los tubos.

3.2.4. Diagrama de cajas o boxplot

Los tres valores de la variable que divide la distribución de frecuencias, previamente ordenada en forma ascendente, en cuatro partes iguales, se denominan *cuartiles*. Note que el porcentaje de valores entre ellos es del 25 %. Usualmente se denotan así:

C_1 : Cuartil 1. Tiene el 25 % de los datos por debajo de él y el 75 % restante por encima.

C_2 : Cuartil 2. Coincide con la mediana. Tiene el 50 % de los datos por debajo y el 50 % restante por encima.

C_3 : Cuartil 3. Tiene el 75 % de los datos por debajo de él y el 25 % restante por encima.

El diagrama de cajas se basa en los cuartiles y sirve para detectar valores atípicos en los datos, así como para estudiar la simetría de la distribución de la variable y la homogeneidad de los datos. Este gráfico es muy útil para comparar varios grupos de datos.

Los pasos a seguir son:

1. Calcule los cuartiles de la distribución.
2. Determine:

$$\text{Rango intercuartílico } Q = C_3 - C_1$$

$$\text{Cercos interior inferior } C_{II} = C_1 - 1.5(Q)$$

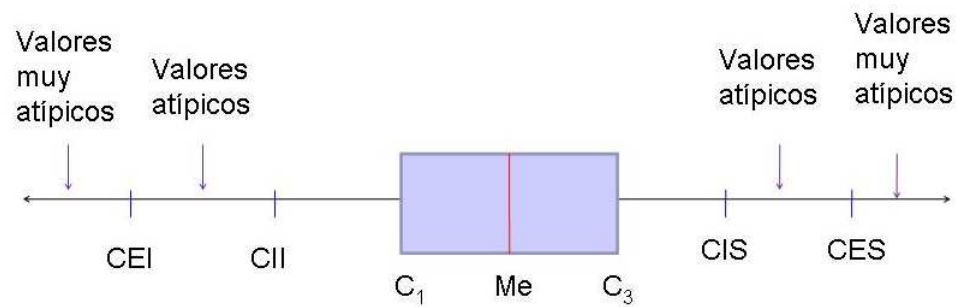
Cerco interior superior $C_{IS} = C_3 + 1.5(Q)$

Cerco exterior inferior $C_{EI} = C_1 - 3.0(Q)$

Cerco exterior superior $C_{ES} = C_3 + 3.0(Q)$

3. Trace una caja de ancho arbitrario, pero con la misma longitud del rango intercuartílico.

Ubique los valores hallados anteriormente, como se muestra en el siguiente gráfico:



4. Finalmente, ubique los valores de la variable en la recta trazada en el gráfico.

Para concluir sobre la simetría de la distribución se tiene en cuenta los siguientes criterios:

- Si la línea que representa la mediana de la distribución está cercana a C_1 , entonces la distribución es asimétrica negativa.
- Si la línea que representa la mediana de la distribución está centrada, entonces la distribución es simétrica.
- Si la línea que representa la mediana de la distribución está cercana a C_3 , entonces la distribución es asimétrica positiva.

La homogeneidad de los datos la determina el ancho de la caja, entre más larga sea la caja más heterogéneos son los datos.

Cuando se comparan varios grupos, además de los anteriores criterios, se considera que si las cajas de los grupos se translapan entonces no existen diferencias significativas en los valores de los grupos.

■ **Ejemplo 12.** Para los diámetros de los tubos tenemos:

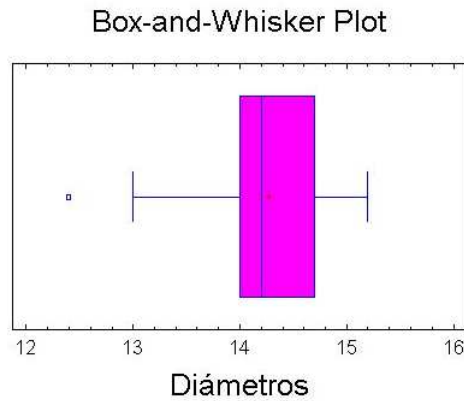


Figura 11: Diagrama de cajas de los diámetros de los tubos.

En la figura 11 se nota que la distribución de los diámetros de los tubos es asimétrica negativa. Así mismo, se observa que los datos son muy homogéneos, lo que indica que la media es una buena representación de los datos. Así mismo, se observa un valor atípico que corresponde a al mínimo valor que toma la variable: 12.4

■ **Ejemplo 13.** En la figura 12 se muestra el diagrama de cajas para comparar los casos mensuales nuevos de malaria en Antioquía por período epidemiológico (1998-2002).

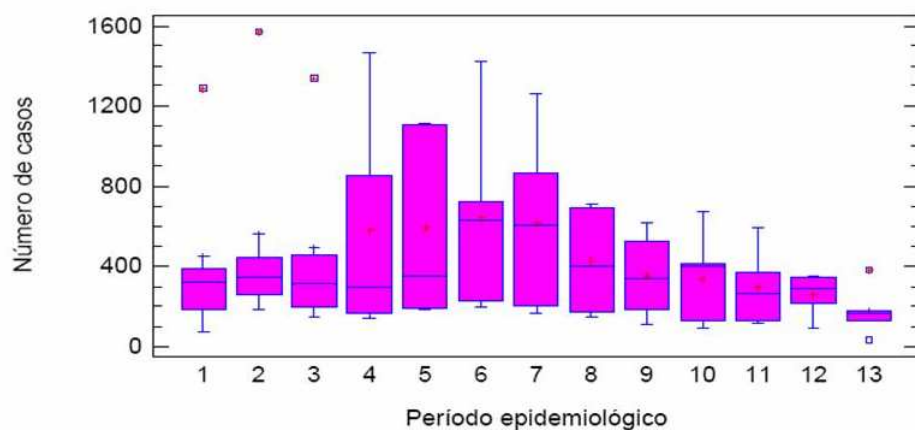


Figura 12: Casos mensuales nuevos de malaria en Antioquía (1998-2002).

A partir de la figura 12 se concluye que en el período 5 se presentó un mayor número de casos de malaria. En general no existen diferencia significativas entre los números de caos de malaria

registrados en los diferentes períodos, excepto el período 13 que difiere de los otros períodos; y se caracteriza por presentar un menor número de casos registrados. Además, se nota que el número de casos registrados en los períodos 4 a 10 fueron muy heterogéneos, a diferencia del 13 que es el más homogéneo. Por otra parte, se nota que en los períodos 4, 5 y 11 fue más frecuente registrar grandes números de casos de malaria, a diferencia de los períodos 6, 7, 9 y 12, en donde fue más frecuente registrar un menor número de casos.

Una limitación del diagrama de cajas es que no permite identificar distribuciones bimodales:

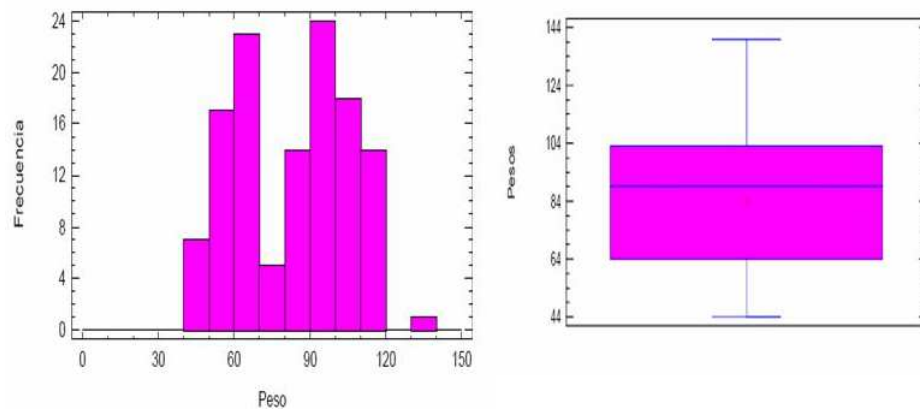


Figura 13: Ilustración de la limitación de los diagramas de cajas.

4. Indicadores

4.1. Razón y/o índice, Tasa y Proporción

En muchos casos es de interés analizar la población y su comportamiento en el tiempo a través de ciertas características. Para hacer esto, se debe cuantificar el comportamiento de la población, midiendo sus características principales a través de relaciones matemáticas como:

- Razones y/o índices.
- Tasas.
- Proporciones.

A continuación se definen cada una de estas relaciones matemáticas, las cuales son necesarias para la comprensión de esta sección:

Razón y/o Índice: Desde el punto de vista de la población, la razón mide la relación entre subgrupos de una población, es decir, un subgrupo dividido por otro. Por ejemplo, suponga que la población de un salón de clases es de 40 estudiantes de los cuales 18 son hombres y el resto mujeres, entonces:

- **Razón de masculinidad:** Mide la relación de hombres versus mujeres.

$$\frac{N^{\circ} \text{ de hombres}}{N^{\circ} \text{ de mujeres}} = \frac{18}{22} = 0.82$$

Por cada mujer hay 0.82 hombres.

- **Razón de femineidad:** Mide la relación de mujeres versus hombres.

$$\frac{N^{\circ} \text{ de mujeres}}{N^{\circ} \text{ de hombres}} = \frac{22}{18} = 1.22$$

Por cada hombre hay 1.22 mujeres.

Cuando la razón entre dos valores se expresa como porcentaje, el valor obtenido se denomina *Índice*. Para el ejemplo tenemos que el índice de masculinidad es $0.82 \times 100 = 82$, es decir, por cada 100 mujeres hay 82 hombres en el salón de clases.

Para interpretar una razón o un índice se debe partir del denominador, y éste a su vez se relaciona con la constante.

Tasas: Se refiere a la frecuencia con que ocurren los eventos poblacionales en un período determinado, es decir, que tan comunes son éstos eventos. Por ejemplo, la Tasa de Natalidad indica con qué frecuencia ocurren los nacimientos. Las tasas están acompañadas de una constante que toma valores generalmente de 1000 en adelante para movimientos poblacionales. Dependiendo del evento demográfico, las tasas se multiplican por 100, 1000, 10000 o 100000.

Proporción: Se refiere a la relación entre un subgrupo de la población y toda la población en un período determinado. Por ejemplo, si de 100 personas 60 son mujeres, entonces el 60 % del total son mujeres. Se acompaña siempre con una constante igual a 100.

4.2. Definición e importancia de un indicador

Un indicador es un instrumento de medición que establece una relación entre dos o más datos significativos, y que proporciona información sobre el estado de eficiencia, eficacia, en que se encuentra un proceso, un área.

Los indicadores son elementos informativos del control de cómo funciona una actividad, pues hacen referencia a parámetros estables que sirven de magnitud de comprobación del funcionamiento de ésta. Son los elementos básicos de las técnicas de control de gestión. La utilidad y fiabilidad del control de gestión se vincula necesariamente a la utilidad y fiabilidad de los indicadores. Una vez elegidos los indicadores, se imponen sobre ellos técnicas de seguimiento : ratios, precios de coste, organigramas, presupuestos. Para ser adecuados, los indicadores deben ser:

- Accesibles (fáciles de identificar y recopilar).
- Pertinentes (para los que deseamos medir).
- Fieles (que informen con fidelidad de las condiciones de los datos que se recogen).
- Objetivos (no ambiguos en su interpretación).
- Precisos (para la acción que se quiere estimar).
- Unívocos (parámetros exclusivos de lo que se mide) y Sensibles (que permitan recoger y estimar variaciones de aquello que son referente).

Es necesario resaltar la diferencia que existe entre un indicador y un índice: La combinación de varios indicadores nos permite construir un índice. Por ejemplo, el Índice de Desarrollo Humano se basa en un indicador social compuesto por tres aspectos:

- Vida larga y saludable (medida según la esperanza de vida al nacer).
- Educación (medida por la tasa de alfabetización de adultos y la tasa bruta combinada de matriculación en educación primaria, secundaria y superior).
- Nivel de vida seguro (medido por el Producto Interno Bruto per cápita en dólares).

4.2.1. Tipos de indicadores

Indicadores de ejecución: Son los que aluden a resultados de la actividad. Pueden ser:

- De economía: Número recursos empleados/disponibles.
- De eficiencia: Valoran los recursos empleados en relación a los resultados concretos obtenidos.
- De eficacia: Comparan los resultados obtenidos con los previstos.
- De efectividad: Valoran el resultado global concreto con el previsto.

Indicadores de proceso: Aluden a los procesos intermedios de la actividad. Pueden ser:

- Estratégicos: Informan de qué factores externos influyen en el proceso de actividad
- De Estructura: Valoran los recursos disponibles y los necesarios.
- De Proceso: Evalúan cómo se desarrollan las actividades intermedias del proceso de gestión
- De Resultado: Miden los resultados finales del proceso

4.3. Pirámides e Indicadores demográficos

4.3.1. Pirámide poblacional

por lo que el estudio de la estructura por edad de una población constituye un elemento básico de la demografía. Muchos fenómenos sociales están influidos, entre otros factores, por la estructura etárea de la población. Para representar la estructura por género y edad se acostumbra a recurrir a la llamada pirámide de población, que no es otra cosa que uno o dos histogramas (según se represente uno o dos géneros), cuyas barras tienen base proporcional a la amplitud del intervalo de edad y superficie proporcional a la población (o porcentaje que ésta representa respecto del total) de los grupos. Cuanto mayor sea el número de grupos de edad, mayor será el detalle con que se describa la estructura poblacional². Una pirámide poblacional encierra información sobre conflictos sociales y guerras, efectos indirectos de éstas como fenómenos migratorios, caídas y recuperaciones de la natalidad, cambios económicos, etc. Se tiene por tanto que la forma y evolución de la pirámide depende directamente de las entradas y salidas a la misma, es decir de los nacimientos y defunciones durante el último siglo aproximadamente (aunque también pueden tener gran influencia los movimientos migratorios). De acuerdo a los nacimientos y defunciones registrados en los pasados 100 años se obtendrá una u otra forma de pirámide poblacional y, aunque serían infinitas las configuraciones posibles, se pueden agrupar

en unas cuantas tipologías que a grandes rasgos permiten caracterizar a todas las poblaciones reales¹. Se pueden considerar 3 grandes tipos estructurales, si bien existen formas intermedias o de transición. El primero de ellos, el denominado *pagoda*, tiene base ancha y disminuye rápidamente hacia el vértice; corresponde a poblaciones jóvenes con alta natalidad y bajo nivel de desarrollo. El tipo *campana* es de base intermedia y disminución lenta hacia el vértice; corresponde a poblaciones estacionarias y que comienzan a envejecer. Por último, el tipo *hucha* es de base estrecha, se ensancha en el centro y disminuye suavemente hacia el vértice; corresponde a poblaciones donde la fecundidad es muy baja y se encuentran en proceso de envejecimiento.

Si una población es estacionaria (igual número de muertes que de nacimientos), cuanto más pequeños sean los flujos de entrada y salida, más envejecida será la estructura de edad y la media de edad de los individuos será mayor. Por el contrario, cuanto más alta sea la mortalidad y la natalidad, la media de edad de los individuos será más baja y, por tanto, la estructura de edad de la población más joven.

La proporción de ancianos en una población depende sólo parcialmente del número de éstos (numerador), pero depende fuertemente del denominador (el tamaño de la población en su conjunto), que a su vez está en función del comportamiento reproductivo, de la mortalidad y de las migraciones a través del tiempo. Es por ello que el envejecimiento de la población que se está produciendo en las sociedades desarrolladas es consecuencia del descenso de la natalidad durante el último siglo, mientras que el descenso de la mortalidad apenas ha influido en este proceso, ya que se produjo en todas las edades.

4.3.2. Indicadores demográficos

Con el objetivo de sintetizar la distribución de la población según género y edad se suele apelar a una serie de indicadores demográficos adaptados a las distintas necesidades de síntesis y comparación. Todos ellos se calculan a partir de los mismos datos necesarios para construir las pirámides y algunos sólo se podrán calcular cuando se construyen éstas con grupos quinquenales de edad. A continuación se describen algunos de los más utilizados:

Edad media: Se estima computando la media ponderada de los valores centrales de las clases usando las frecuencias relativas de cada grupo como factores de ponderación.

Edad mediana: Es aquella que divide en dos partes iguales a la población cuyos habitantes han sido ordenados según la edad. Es más variable que la edad media pero es menos sensible

a los cambios estructurales propios de cada población.

Índice de envejecimiento (IV): Indicador sintético del grado de envejecimiento de la población; se obtiene dividiendo el conjunto de la población anciana, a partir de los 65 años, entre el de los niños por debajo de cierta edad, generalmente los 15 años.

Índice demográfico de dependencia (ID): Tiene relevancia económica y social. Con este indicador, las personas que supuestamente no son autónomas por razones demográficas (la edad), es decir, los ancianos (> 65) y los muy jóvenes (< 15), se relacionan con las personas que supuestamente deben sostenerlas con su actividad (15-64).

Índice de estructura de la población activa (IS): Es un indicador del grado de envejecimiento de este sector de la población. Puede obtenerse dividiendo la población entre los 40 y los 64 años (las 25 generaciones más viejas en activo) por la población desde los 15 a los 39 años (las 25 generaciones más jóvenes). Cuanto más bajo sea el índice, más joven es la estructura de la población laboral.

Índice de reemplazamiento de la población en edad activa (IR): Es el cociente formado por los que están a punto de salir de la edad activa (60-64 años) entre los que están a punto de entrar (15-19 años). Es un índice sujeto a fuertes fluctuaciones y por tanto muy variable¹.

Índice del número de niños por mujer fecunda (IC): Se obtiene dividiendo los niños nacidos recientemente (0-4 años) por las mujeres en edad fértil (15-49 años). El IC en los países subdesarrollados no es un buen indicador de la fecundidad debido a la merma que supone la elevada mortalidad infantil. Realmente es un indicador de la carga de hijos en edad preescolar por mujer, por lo que es útil como indicador socio-demográfico.

Tasa general de fecundidad (GF): Se obtiene del cociente entre los nacidos vivos durante un año y la población femenina en edad fértil (15-49 años) y expresa el número de hijos que en promedio tendría una cohorte ficticia de mujeres no expuestas al riesgo de morir desde el nacimiento hasta el término de la edad fértil.

Índice de masculinidad (IM): Es el número de nacimientos masculinos por cada 100 nacimientos femeninos. Este índice también se calcula para cada edad.

Índice de Friz (IF): Representa la proporción de población en el grupo 0-19 años en relación a la de 30-49 años, que se toma como base 100. Si este índice es mayor de 160, la población se considera joven, si se halla entre 60 y 160, madura, y si es menor de 60, vieja

Índice de Sundbarg (IS): Tomando como base la población de 15-49 años, Sundbarg compara gráficamente los porcentajes que, en relación con este grupo (= 100), significan los grupos de 0-14 años y de mayores de 50. Si el porcentaje de los menores de 15 supera el de los mayores de 50 la población es de carácter progresivo.

4.4. Indicadores de calidad de vida

4.4.1. Indicadores de Impacto

Necesidades Básicas Insatisfechas (NBI): Las Necesidades Básicas Insatisfechas NBI, se definen como las carencias que tienen los hogares en ciertos bienes y servicios que se consideran básicos para su subsistencia. Las NBI están definidas por los siguientes cinco indicadores:

1. Viviendas inadecuadas.
2. Hacinamiento crítico.
3. Servicios inadecuados.
4. Alta dependencia económica.
5. Inasistencia escolar.

Puesto que cada indicador simple se refiere a una necesidad básica diferente, se construye un indicador compuesto a partir de ellos, que clasifica como pobre o con NBI, a toda vivienda que se encuentre en al menos una de las situaciones de carencia expresada por los indicadores simples, y en miseria toda vivienda que tenga dos o más de los indicadores simples de necesidades básicas insatisfechas.

Para estimar la magnitud de la pobreza en la población, se asume que los hogares y personas que habitan en viviendas con NBI o miseria, se encuentran en esas mismas condiciones. La magnitud de la pobreza se calcula como el total de hogares que poseen un indicador sobre el total de hogares.

4.4.2. Indicadores sectoriales

Estudiaremos dos aspectos: Fuerza laboral y la educación.

1. **Fuerza laboral:** En Colombia los estudios de fuerza de trabajo, tradicionalmente consideran dentro del estudio del Mercado Laboral el concepto de ocupados (PO) como las personas que en la semana de referencia:

- trabajaron la mayor parte del tiempo,
- trabajaron por lo menos una hora remunerada,
- no trabajaron pero tenían trabajo, o
- los trabajadores sin remuneración que trabajaron por lo menos una hora.
- Además define a los desempleados o población desocupada (PD) como las personas que la semana de referencia no tenían empleo la semana de referencia:
- buscaron trabajo la mayor parte del tiempo
- hicieron diligencias el ultimo mes para encontrar uno y además estaban disponibles para trabajar, o
- no hicieron diligencias en el ultimo mes, pero si en los últimos 12 meses y estaban disponibles para empezar a trabajar.

Así la suma entre los Ocupados y los Desocupados constituye la Población Económicamente Activa (PEA).

Con base en esta aclaración se definen las siguientes tasas:

- **Tasa de Ocupación (TO):** se define como el cociente entre la población ocupada (PO) y la población en edad de trabajar (PET). En este caso, PET se define como la población de 10 años y más.
- **Tasa de Desempleo (TE):** Se define como el cociente entre la población desocupada (PD) y la población económicamente activa (PEA). Ésta última la conforman las personas en edad de trabajar que trabajan (ocupados) o desean trabajar (desocupados).
- **Tasa Global de Participación (TGP):** esta dada por el cociente entre la población económicamente activa (PEA) y la población en edad de trabajar (PET). Puesto que la PEA es un subconjunto de la PET, este indicador varía entre 0 y 100. Como

consecuencia de la forma de medición de los ocupados y los desocupados, se espera que los inactivos aumenten y la TGP se subestime.

2. **Educación:** Se describen a continuación las tasas de asistencia escolar por rangos de edad, y, de cobertura bruta y neta por nivel educativo: preescolar, primaria, secundaria y superior.

- **Tasa de asistencia escolar (TAE):** Se define como la proporción de la población en un rango de edad determinado que asiste a un centro de educación formal.
- **Tasa de cobertura bruta (TCB):** Corresponde al cociente entre el número de alumnos matriculados en un nivel educativo (independiente de la edad que tengan) y la población escolar que tiene la edad apropiada para estar en ese nivel.

Los rangos de edad apropiada para cursar un nivel determinado son:

- En preescolar deben estar las personas entre 5 y 6 años
- En primaria deben estar las personas entre 7 y 11 años
- En secundaria deben estar las personas entre 12 y 17 años
- En superior deben estar las personas entre 18 y 24 años

Este indicador puede tomar valores por encima de 100 debido a la extraedad, esto es, la presencia de alumnos cursando niveles con edades superiores a la edad apropiada.

- **Tasa de cobertura neta (TCN):** Es la razón entre el número de alumnos matriculados en un nivel educativo que tienen la edad adecuada para cursarlo y la población escolar que tiene la edad apropiada para estar en ese nivel.
- **Tasa de cobertura neta ajustada por nivel (TCNANIV):** Indica el porcentaje total de niños matriculados por nivel educativo que están en edad escolar (5 a 17 años) y cuentan con la edad apropiada definida para cada nivel educativo.

Bibliografía

- [1] Martinez C.(2007). *Estadística Básica Aplicada*. Tercera Edición. ECOE Ediciones.
- [2] Bautista L.(1998). *Diseños de muestreo estadístico*. Universidad Nacional de Colombia.
- [3] Portocarrero J.(2007). *La medición de indicadores en la investigación*. Editorial USC.
- [4] Lora E.(2007). *Técnicas de medición económica*. Cuarta Edición. Editorial Alfaomega.
- [5] Salvador M. and Gargallo P.(2003). *Análisis Exploratorio de Datos*. [en línea] 5campus.com, Estadística <http://www.5campus.com/leccion/aed>.